Supporting Information

Hansen et al. 10.1073/pnas.0912668107

SI Methods

MODIS Forest Cover Loss Indicator Mapping. The humid tropical, dry tropical, temperate, and boreal biomes were delineated by using the World Wildlife Fund ecoregions map (1) as the main reference. Biome-wide forest cover loss indicator maps were created by using annual MODIS imagery for 2000–2005. Classification and regression tree bagging algorithms (2, 3) were applied to generate per pixel annual and 4- and 5-year forest cover loss probability and forest cover loss proportion maps. Forest cover loss probability maps were made for the dry tropical and humid tropical biomes, and forest cover loss proportion maps were made for the temperate and boreal biomes. MODIS 32-day composites were used as inputs and included all land bands: blue (459–479 nm), green (545–565 nm), red (620–670 nm), near infrared (841–876 nm), and mid-infrared (1230–1250, 1628–1652, 2105–2155 nm) (4). In addition, data from the MODIS Land Surface Temperature product (5) were also used as inputs. The time-sequential MODIS 32-day inputs were transformed to annual metrics to produce a more generalized feature space. Annual image metrics have been shown to perform as well or better than time-sequential composites for mapping large areas (6). The supervised tree algorithms related the expert-interpreted training data to the MODIS metrics. For each MODIS forest cover loss map, the per-pixel probability of forest cover loss or proportion of forest cover loss was aggregated to the 18.5-km × 18.5-km block scale. MODIS forest cover loss probability maps for the humid and dry tropical biomes were thresholded into discrete loss/no loss categories to calculate percent of forest cover loss per block. MODIS maps of fraction of forest cover loss for the temperate and boreal biomes were averaged to the block scale. Percent indicated forest cover loss per block was used in the stratification and regression estimation procedures.

Landsat Sample Block Forest Cover and Loss Mapping. Per sample block analyses were performed by using Landsat ETM+ image pairs for a total of 541 sample blocks (Fig. S1). Forest cover 2000 and gross forest cover loss (GFCL) 2000–2005 were classified by using a supervised decision tree classifier (2). Each block was examined in detail by one or more interpreters and the classification process iterated, if needed. All tree cover assemblages that met the 25% canopy closure definition used for this study, whether intact forests, plantations, or forest regrowth, were defined as forests. Missing data per sample block included hand-delineated cloud and shadow cover and data gaps from the Landsat 7 ETM+ Scan Line Corrector-Off (SLC-off) malfunction. The Landsat forest cover and forest loss data for the 541 sample blocks may be accessed at http://globalmonitoring.sdstate.edu/projects/gfm/global/gindex.html. The estimated mean GFCL per stratum within each biome derived from the Landsat imagery is provided in Table S1.

Sampling. The sampling strategy implemented employed stratified random sampling combined with a separate regression estimator (i.e., the regression relationship is established separately for each stratum) (7, 8). The stratification was determined from the MODIS-forest change products. Stratum breakpoints for each biome were initially determined by applying the Dalenius-Hodges rule (7), but modified slightly to reduce the size of the low-change strata under the assumption that the MODIS data would tend to underestimate actual GFCL. A retrospective evaluation of the MODIS-based stratification revealed that the strata selected were highly effective (Table S2). The GFCL data derived from the Landsat interpretation of the sample blocks can be used to estimate the standard errors that would have been obtained had stratification not been implemented (i.e., if instead simple random sampling had been used). For the five major regions sampled (the four biomes with the humid tropics split into two major strata, Indonesia and outside Indonesia), stratification improved upon the precision of simple random sampling. Except for the humid tropics outside of Indonesia, the improvement in standard error achieved by the stratified design was substantial. These results demonstrate the effectiveness of using the MODIS forest cover loss data to provide a spatially fine-grained stratification of the 18.5-km × 18.5-km blocks. This highly targeted spatial stratification offers an improvement over a more generalized “hot spot” stratification in which much larger areas in a more subjective fashion are delineated to define low and high forest clearing strata (9).

Regression Estimators. Survey sampling regression estimators (7, 8) exploit ancillary variables related to the target variable of interest to improve precision of estimates. The ancillary variables used to improve precision of the estimates of GFCL included several variables constructed from the MODIS forest cover loss indicator maps. Regression estimators were not used in the lowest change stratum of each biome because the relationship between the Landsat-determined GFCL and MODIS-determined GFCL was too weak to provide a useful reduction in the standard error. Poststratified estimation (8) was employed in several of the low change strata to improve precision because it was possible to effectively subdivide the low change strata into poststrata representing virtually no change and some change. The ancillary variables used to define the poststrata included percent tree cover (10) and area of intact forest landscape (11).

Sampling of Satellite Imagery. Debate on the proper use of Landsat data for regional and global monitoring has concerned exhaustive mapping versus sample-based approaches (12–14). Data limitations, primarily cloud cover and costs of imagery, have limited exhaustive mapping exercises. The difficulty with the sampling approach is that forest cover change is typically rare, and perhaps strongly clustered, at regional, biome, or global scales, and simple random sampling of Landsat scenes has been demonstrated to be inadequate for country specific estimates in some cases (12). The suggestion to use a sampling unit smaller than a Landsat scene (12) has been shown to lead to reduced standard errors of area estimates of change (15). Two other features of the sampling strategy achieved substantial improvements in the precision of the GFCL area estimates. The stratification based on the MODIS-derived forest loss effectively partitioned each biome into regions of relatively homogeneous GFCL, as demonstrated by the reduction in standard error relative to simple random sampling (an unstratified design) (Table S2). Incorporating ancillary variables via poststratified and regression estimators produced further reductions in the standard errors of the area estimates. The ultimate measure of whether the difficulties of a sampling-based approach (12) have been overcome is the standard error of the area estimate. The standard errors presented in Tables 1–3 are generally sufficiently low to indicate that the combination of MODIS-based stratification with poststratified and regression estimators resulted in precise estimates of area of GFCL.

Comparison with INPE’s PRODES Data. PRODES data from 2000 to 2005 were compared with the humid tropical block-calibrated results on a per state basis for the Legal Amazon. PRODES deforestation pixel counts were totaled as a fraction of year 2000 PRODES forest area within 18.5-km × 18.5-km sample blocks where PRODES 2000 forest extent exceeded 50% of the block area. Fig. S5 illustrates the correspondence of PRODES and MODIS/Landsat percent forest loss per state. Area totals for this analysis do not equal official PRODES estimates because the blocks do not exactly match the PRODES study extents and PRODES official results are not based solely on pixel counts. However, there is strong agreement between the two remotely sensed forest loss estimation methods. The comparison with PRODES is meant for evaluation purposes only. The precision (standard error) of the GFCL estimate for any subregion is a function of sample size. The results presented in Tables 1–3 include subregions for which reasonably precise estimates were obtained to support meaningful comparisons among subregions.

1. Olson D, et al. (2001) Terrestrial ecoregions of the World: A new map of life on Earth. *Bioscience* 51:1–6.
2. Breiman L, Friedman J, Olshen R, Stone C (1984) *Classification and Regression Trees* (Wadsworth and Brooks/Cole, Monterey, CA).
3. Breiman L (1996) Bagging predictors. *Mach Learn* 26:123–140.
4. Wolfe R, Roy D, Vermote E (1998) MODIS land data storage, Gridding, and compositing methodology: Level 2 grid. *IEEE Trans Geosci Rem Sens* 36:1324–1338.
5. Wan Z, Zhang Y, Zhang Q, Li Z-L (2002) Validation of the land surface temperature products retrieved from Terra Moderate Resolution Imaging Spectroradiometer data. *Remote Sens Environ* 83:163–180.
6. Hansen M, Townshend J, DeFries R, Carroll M (2005) Estimation of tree cover using MODIS data at global, continental and regional/local scales. *Int J Remote Sens* 26: 4359–4380.
7. Cochran W (1977) *Sampling Techniques* (Wiley, New York), 3rd Ed.
8. Sarndal C-E, Swensson B, Wretman J (1992) *Model-Assisted Survey Sampling* (Springer, New York).
9. Achard F, et al. (2002) Determination of deforestation rates of the world's humid tropical forests. *Science* 297:999–1002.
10. Hansen M, et al. (2003) Global percent tree cover at a spatial resolution of 500 meters: First results of the MODIS Vegetation Continuous Fields algorithm. *Earth Interact* 7: 10.1175/1087-3562.
11. Potapov P, et al. (2008) Mapping the world's intact forest landscapes by remote sensing. *Ecol Soc* 13:51.
12. Tucker CJ, Townshend JRG (2000) Strategies for monitoring tropical deforestation using satellite data. *Int J Remote Sens* 21:1461–1471.
13. Grainger A (2008) Difficulties in tracking the long-term global trend in tropical forest area. *Proc Natl Acad Sci USA* 105:818–823.
14. Stehman SV (2005) Comparing estimators of gross change derived from complete coverage mapping versus statistical sampling of remotely sensed data. *Remote Sens Environ* 96:466–474.
15. Broich M, et al. (2009) A comparison of sampling designs for estimating deforestation from Landsat imagery: A case study of the Brazilian Legal Amazon. *Remote Sens Environ* 113:2448–2454.

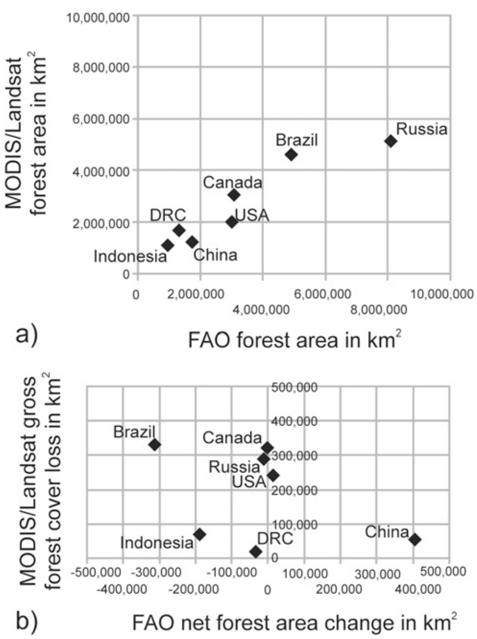

Fig. S1. Comparison of FAO and MODIS/Landsat estimates for 2000 forest area (*A*) and 2000–2005 FAO net forest area change and MODIS/Landsat gross forest cover loss (GFCL) (*B*).

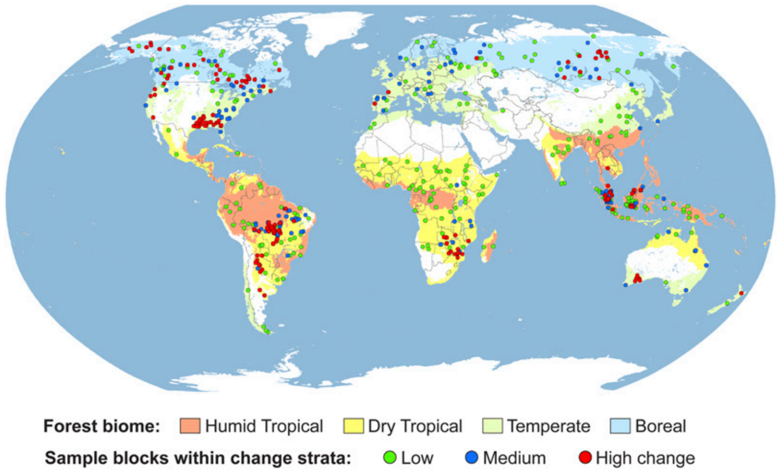

Fig. S2. Sample block locations by biome and MODIS-indicated gross forest cover loss stratum.

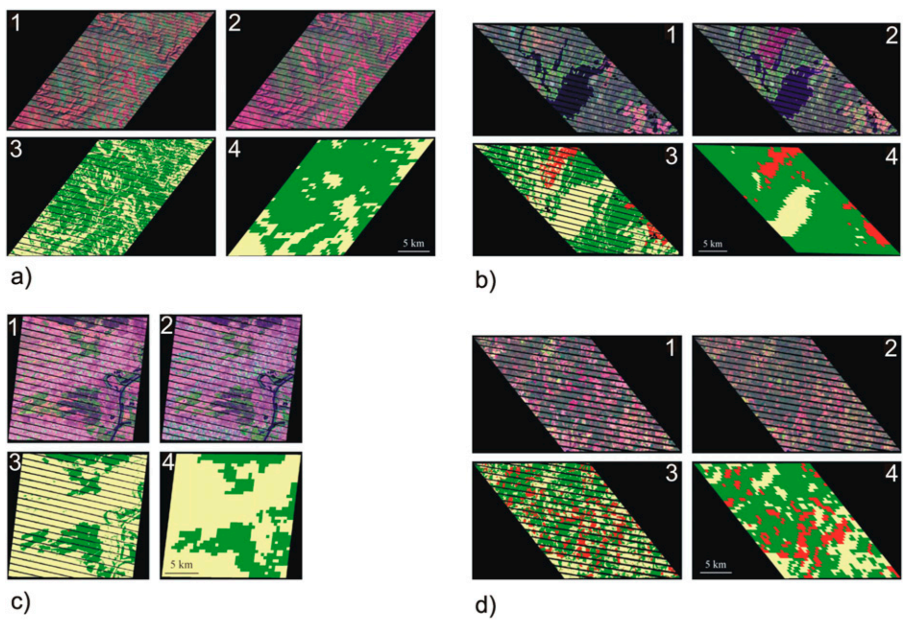

Fig. S3. (*A*) Example boreal forest biome sample block, low MODIS-indicated change stratum. Location: 52.1° N; 57.6° E; Europe, Russia, Bashkortostan Republic, Landsat: WRS Path 164 Row 24; Image date (1): 9/23/2001 (2), 8/17/2005 Landsat results (3): Data 89.3%, Forest: 58.6%, Change 0.2% MODIS results (4): Hotspots 0.0%; Mean VCF 31%. (*B*) Example boreal forest biome sample block, high MODIS-indicated change stratum. Location: 48.8° N; 77° W; North America, Canada, Quebec, Landsat: WRS Path 17 Row 26; Image date (1): 6/13/2001 (2), 12/9/2005 Landsat results (3): Data 79.9%, Forest: 54.6%, Change 6.6% (Burned 4.2%) MODIS results (4): Hotspots 9.0% (Burned 4.13%); Mean VCF 52%. (*C*) Example temperate forest biome sample block, low MODIS-indicated change stratum. Location: 49.3° N; 8.3° E; Europe, Germany, Rheinland-Pfalz Landsat: WRS Path 195 Row 26; Image date (1): 9/11/1999 (2), 6/23/2005 Landsat results (3): Data 85.8%, Forest: 24.4%, Change 0.1% MODIS results (4): Hotspots 0.2%; Mean VCF 22%. (*D*) Example temperate forest biome sample block, high MODIS-indicated change stratum. Location: 31.1° N; 86.3° W; North America, United States, Alabama Landsat: WRS Path 20 Row 38; Image date (1): 9/17/1999 (2), 9/1/2005 Landsat results (3): Data 74.9%, Forest: 62.0%, Change 15.2% MODIS results (4): Hotspots 4.9%; Mean VCF 35%.

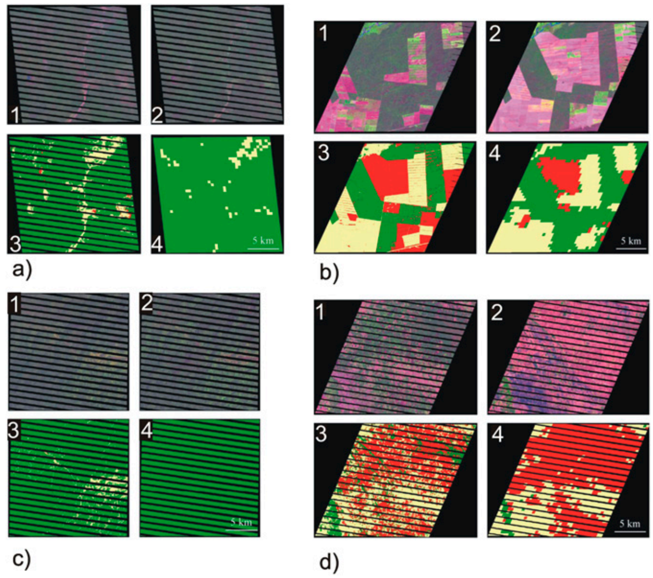

Fig. S4. (*A*) Example dry tropical forest biome sample block, low MODIS-indicated change stratum. Location: 13.8° S; 24.6° E; Africa, Zambia, North-Western Landsat: WRS Path 174 Row 70; Image date (1): 5/8/2001 (2), 4/17/2005 Landsat results (3): Data 71.6%, Forest: 89.7%, Change 0.7% MODIS results (4): Stratum 1; Hotspots 0.0%; Mean VCF 40%. (*B*) Example dry tropical forest biome sample block, high MODIS-indicated change stratum. Location: 24.6° S; 63.9° W; South America, Argentina, Salta, Landsat: WRS Path 230 Row 77; Image date (1): 4/14/2001 (2), 1/6/2006 Landsat results (3): Data 99.7%, Forest: 62.7%, Change 26.1% MODIS results (4): Stratum 3; Hotspots 12.0%; Mean VCF 30%. (*C*) Example humid tropical forest biome sample block, low MODIS-indicated change stratum. Location: 3.42° N; 15.61° E; Africa, Central African Republic, Sangha, Landsat: WRS Path 183 Row 58; Image date (1): 12/14/2000 (2), 1/10/2005 Landsat results (3): Data 65.2%, Forest: 92.9%, Change 0.1% MODIS results (4): Stratum 1; Hotspots 0.0%; Mean VCF 72%. (*D*) Example humid tropical forest biome sample block, high MODIS-indicated change stratum. Location: 13.25° N; 102.48° E; Asia, Cambodia, Batdambang, Landsat: WRS Path 128 Row 51; Image date (1): 11/30/2001 (2), 2/26/2005 Landsat results (3): Data 76.1%, Forest: 57.8%, Change 44.1% MODIS results (4): Stratum 4; Hotspots 65.0%; Mean VCF 42%.

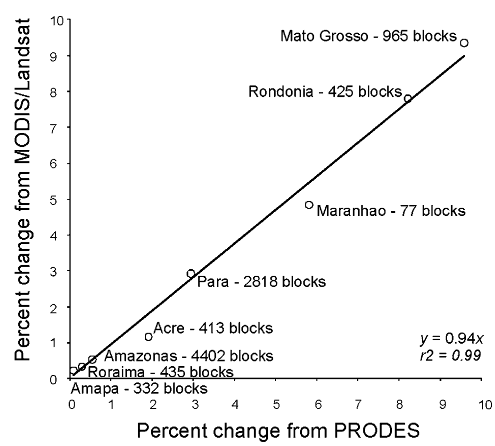

Fig. S5. Comparison of percent forest change for regions common to both the MODIS/Landsat humid tropical biome GFCL and INPE'S PRODES deforestation data sets for 2000–2005.

Table S1. Landsat-derived estimated GFCL by stratum

Stratum	MODIS-derived GFCL, %	Landsat-derived GFCL, %	No. of blocks sampled
Humid Tropics—Indonesia			
1	0–2%	1.39	41
2	2–9%	11.79	17
3	>9%	25.92	18
4	"certainty"	72.35	1
Humid tropics—outside of Indonesia			
1	0–2	1.03	46
2	2–9	11.48	23
3	>9	21.25	32
4	"certainty"	38.45	5
Boreal			
1	0–0.25	0.50	25
2	0.25–1	1.23	25
3	1–6	4.53	25
4	>6%	13.13	43
Dry tropics			
1	0	0.51	65
2	0–1	4.30	25
3	>1	6.88	30
Temperate			
1	0–0.25	0.25	40
2	0.25–3	1.78	40
3	>3	12.35	40

The sample means for Landsat-derived GFCL are based solely on the Landsat data and are not the stratum-specific means that would be obtained from a regression or poststratified estimator.

Table S2. Retrospective evaluation of effectiveness of MODIS-based stratification

Region	SE ratio	Stratified sample size	Equivalent SRS sample size
Temperate	2.89	120	1002
Boreal	2.07	120	514
Dry Tropics	1.51	120	274
Humid Tropics sans Indonesia	1.16	72	97
Indonesia (Humid Tropics)	2.04	76	316

The SE Ratio is the SE of the estimated area of GFCL for simple random sampling (SRS) divided by the SE for the stratified design implemented. The Equivalent SRS Sample Size is the number of sample blocks that would be required when using simple random sampling to achieve the same standard error for estimated area of GFCL as was obtained from the MODIS-based stratified design.